

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:38

INPUT SET: S34959.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information
4
5 (i) APPLICANT: Murphy, Dennis
6 Reid, John
7
8 (ii) TITLE OF THE INVENTION: ALPHA-GALACTOSIDASE
9
10 (iii) NUMBER OF SEQUENCES: 4
11
12 (iv) CORRESPONDENCE ADDRESS:
13 (A) ADDRESSEE: Fish & Richardson, P.C.
14 (B) STREET: 4225 Executive Square, Suite 1400
15 (C) CITY: La Jolla
16 (D) STATE: CA
17 (E) COUNTRY: US
18 (F) ZIP: 92037
19
20 (v) COMPUTER READABLE FORM:
21 (A) MEDIUM TYPE: Diskette
22 (B) COMPUTER: IBM Compatible
23 (C) OPERATING SYSTEM: Windows95
24 (D) SOFTWARE: FastSEQ for Windows Version 2.0
25
26 (vi) CURRENT APPLICATION DATA:
27 (A) APPLICATION NUMBER: 09/407,806
28 (B) FILING DATE:
29 (C) CLASSIFICATION:
30
31 (vii) PRIOR APPLICATION DATA:
32 (A) APPLICATION NUMBER: 08/613,220
33 (B) FILING DATE:
34
35 (viii) ATTORNEY/AGENT INFORMATION:
36 (A) NAME: Haile, Ph.D., Lisa A.
37 (B) REGISTRATION NUMBER: 38,347
38 (C) REFERENCE/DOCKET NUMBER: 09010/004001
39
40 (ix) TELECOMMUNICATION INFORMATION:
41 (A) TELEPHONE: 619-678-5070
42 (B) TELEFAX: 619-68-5099
43 (C) TELEX:
44
45 (2) INFORMATION FOR SEQ ID NO:1:
46

RECEIVED
MAR 10 2000
TC 1600 MAIL ROOM

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/407,806ADATE: 03/07/2000
TIME: 10:52:38

INPUT SET: S34959.raw

47 (i) SEQUENCE CHARACTERISTICS:
48 (A) LENGTH: 52 base pairs
49 (B) TYPE: nucleic acid
50 (C) STRANDEDNESS: single
51 (D) TOPOLOGY: linear
52
53 (ii) MOLECULE TYPE: cDNA
54
55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
56
57 CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGAGCG CTCGTCTTTC AC 52
58
59 (2) INFORMATION FOR SEQ ID NO:2:
60
61 (i) SEQUENCE CHARACTERISTICS:
62 (A) LENGTH: 31 base pairs
63 (B) TYPE: nucleic acid
64 (C) STRANDEDNESS: single
65 (D) TOPOLOGY: linear
66
67 (ii) MOLECULE TYPE: cDNA
68
69 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
70
71 CGGAAGATCT AGGTTCCCCA TTTTCACCCC T 31
72
73 (2) INFORMATION FOR SEQ ID NO:3:
74
75 (i) SEQUENCE CHARACTERISTICS:
76 (A) LENGTH: 1041 base pairs
77 (B) TYPE: nucleic acid
78 (C) STRANDEDNESS: single
79 (D) TOPOLOGY: linear
80
81 (ix) FEATURE:
82 (A) NAME/KEY: Coding Sequence
83 (B) LOCATION: 1...1038
84 (D) OTHER INFORMATION:
85
86 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
87
88 TTG AGA GCG CTC GTC TTT CAC GGC AAC CTC CAG TAT GCC GAA ATC CCA 48
89 Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro
90 1 5 10 15
91
92 AAG AGC GAA CCA AAG GTC ATA GAG AAG GCA TAC ATC CCA GTC ATC GAG 96
93 Lys Ser Glu Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile Glu
94 20 25 30
95
96 ACA CTG ATT AAA GAA GAA CCT TTT GGG CTC AAC ATA ACG GGC TAT ACC
97 Thr Leu Ile Lys Glu Glu Pro Phe Gly Leu Asn Ile Thr Gly Tyr Thr
98 35 40 45
99

RECEIVED
MAR 10 2000
TC 1500 MAIL ROOM

RAW SEQUENCE LISTING PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:38

INPUT SET: S34959.raw

100	TTA AAG TTC CTC CCG AAG GAT ATT ATA CTC GTT AAA GGG GGC ATC GCG	192
101	Leu Lys Phe Leu Pro Lys Asp Ile Ile Leu Val Lys Gly Gly Ile Ala	
102	50 55 60	
103		
104	AGT GAC CTG ATA GAG ATA ATC GGA ACG AGC TAC ACG GCA ATA CTC CCC	240
105	Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr Ala Ile Leu Pro	
106	65 70 75 80	
107		
108	CTC CTG CCG CTT AGC AGA GTA GAA GCA CAA GTT CAG AGA GAT AGG GTT	288
109	Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln Arg Asp Arg Val	
110	85 90 95	
111		
112	AAG GAA GAG CTC TTC GAG GTT TCT CCA AAG GGA TTC TGG CTG CCA GAG	336
113	Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly Phe Trp Leu Pro Glu	
114	100 105 110	
115		
116	CTC GCC GAC CCG ATA ATC CCT GCC ATA CTG AAG GAC AAC GGT TAT GAG	384
117	Leu Ala Asp Pro Ile Ile Pro Ala Ile Leu Lys Asp Asn Gly Tyr Glu	
118	115 120 125	
119		
120	TAT CTA TTC GCC GAC GAG GCG ATG CTT TTC TCA GCT CAT CTC AAC TCG	432
121	Tyr Leu Phe Ala Asp Glu Ala Met Leu Phe Ser Ala His Leu Asn Ser	
122	130 135 140	
123		
124	GCG ATA AAG CCA ATT AAA CCG CTC CCA CAC CTT ATA AAG GCC CAA AGG	480
125	Ala Ile Lys Pro Ile Lys Pro Leu Pro His Leu Ile Lys Ala Gln Arg	
126	145 150 155 160	
127		
128	GAA AAG CGC TTT AGG TAC ATC AGC TAT CTC CTT CTC AGG GAG CTT AGG	528
129	Glu Lys Arg Phe Arg Tyr Ile Ser Tyr Leu Leu Leu Arg Glu Leu Arg	
130	165 170 175	
131		
132	AAG GCG ATA AAG CTC GTT TTT GAA GGT AAG GTA ACG CTA AAG GTC AAA	576
133	Lys Ala Ile Lys Leu Val Phe Glu Gly Lys Val Thr Leu Lys Val Lys	
134	180 185 190	
135		
136	GAC ATC GAA GCC GTA CCC GTT TGG GTG GCC GTG AAC ACG GCT GTA ATG	624
137	Asp Ile Glu Ala Val Pro Val Trp Val Ala Val Asn Thr Ala Val Met	
138	195 200 205	
139		
140	CTC ATC GGA AGG CTT CCT CTT ATG AAT CCT AAG AAA GTG GCG AGC TGG	672
141	Leu Ile Gly Arg Leu Pro Leu Met Asn Pro Lys Lys Val Ala Ser Trp	
142	210 215 220	
143		
144	ATA GAG GAC AAG AAC ATT CTT CTA TAC GGC ACC GAT ATA GAG TTC ATT	720
145	Ile Glu Asp Lys Asn Ile Leu Leu Tyr Gly Thr Asp Ile Glu Phe Ile	
146	225 230 235 240	
147		
148	GGC TAT AGG GAC ATT GCA GGC AGA ATG AGT GTT GAG GGA TTA TTA GAG	768
149	Gly Tyr Arg Asp Ile Ala Gly Arg Met Ser Val Glu Gly Leu Leu Glu	
150	245 250 255	
151		
152	GTT ATA GAC GAG CTC AAC TCG GAA CTG TGC CCC TCA GAG CTG AAG CAC	816

RAW SEQUENCE LISTING PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:39

INPUT SET: S34959.raw

```

153 Val Ile Asp Glu Leu Asn Ser Glu Leu Cys Pro Ser Glu Leu Lys His
154          260          265          270
155
156 AGT GGA AGG GAG CTC TAC TTA CGG ACT TCG AGT TGG GCA GAT AAG AGC      864
157 Ser Gly Arg Glu Leu Tyr Leu Arg Thr Ser Ser Trp Ala Asp Lys Ser
158          275          280          285
159
160 TTG AGG ATA TGG AGA GAG GAC GAA GGG AAC GCA AGA CTT AAT ATG CTG      912
161 Leu Arg Ile Trp Arg Glu Asp Glu Gly Asn Ala Arg Leu Asn Met Leu
162          290          295          300
163
164 TAC AAT ATG AGG GGC GAA CTC GCC TTT TTA GCC GAG AAC AGC GAT GCA      960
165 Tyr Asn Met Arg Gly Glu Leu Ala Phe Leu Ala Glu Asn Ser Asp Ala
166          305          310          315          320
167
168 AGG GGA TGG CCC CTC CCT GAG AGG AGG CTG GAT GCC TTC CGG GCG ATA      1008
169 Arg Gly Trp Pro Leu Pro Glu Arg Arg Leu Asp Ala Phe Arg Ala Ile
170          325          330          335
171
172 TAT AAC GAT TGG AGG GGT AAT GGG GAA CCT TAG      1041
173 Tyr Asn Asp Trp Arg Gly Asn Gly Glu Pro
174          340          345
175
176
177 (2) INFORMATION FOR SEQ ID NO:4:
178
179 (i) SEQUENCE CHARACTERISTICS:
180 (A) LENGTH: 346 amino acids
181 (B) TYPE: amino acid
182 (D) TOPOLOGY: linear
183
184 (ii) MOLECULE TYPE: protein
185
186 (v) FRAGMENT TYPE: internal
187
188 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
189
190 Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro
191 1 5 10 15
192 Lys Ser Glu Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile Glu
193 20 25 30
194 Thr Leu Ile Lys Glu Glu Pro Phe Gly Leu Asn Ile Thr Gly Tyr Thr
195 35 40 45
196 Leu Lys Phe Leu Pro Lys Asp Ile Ile Leu Val Lys Gly Gly Ile Ala
197 50 55 60
198 Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr Ala Ile Leu Pro
199 65 70 75 80
200 Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln Arg Asp Arg Val
201 85 90 95
202 Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly Phe Trp Leu Pro Glu
203 100 105 110
204 Leu Ala Asp Pro Ile Ile Pro Ala Ile Leu Lys Asp Asn Gly Tyr Glu
205 115 120 125

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:39

INPUT SET: S34959.raw[illegible]

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:39

INPUT SET: S34959.raw

Line

Error

Original Text

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:39

INPUT SET: S34959.raw

< < THERE ARE NO ITEMS MISSING > >

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/09/407,806A

DATE: 03/07/2000
TIME: 10:52:40

INPUT SET: S34959.raw

Line

Original Text

Corrected Text

(1) General Information

(1) GENERAL INFORMATION:

(ii) TITLE OF THE INVENTION: ALPHA-GALACTOSID (ii) TITLE OF INVENTION: ALPHA-GALACTOSIDASE